Intelligent Data Processing and Mining of Histopathological Images using Improved Tunicate Swarm Algorithm with Deep Learning

Rama Asad Nadweh*[1], Arwa Hajjari[2]

[1] Online Islamic University, Department Of Science and Information Technology, Doha, Qatar
[2] Cairo University, Cairo, Egypt
Emails: ramaanadwehh@gmail.com ; Hajjarint8843@gmail.com

Abstract

Intelligent data processing and mining of histopathological images involve the application of advanced techniques and algorithms to analyze and extract meaningful information from digital pathology images. Osteosarcoma is a general malignant bone cancer generally established in teenagers and children. Manual diagnoses of osteosarcoma is a laborious task and needs skilled professionals. The mortality rate can be minimalized only if it is identified on time. Automatic detection systems and new technologies were utilized to classify and analyze medical images that, minimalize the dependency on specialists and result in fast processing. Recently, a lot of Computer-Aided Diagnosis (CAD) systems were proposed by research workers to diagnose and segment osteosarcoma from medical images. Deep learning (DL) algorithms are employed for the automated recognition and identification of osteosarcoma on histopathological images (HSI). The study proposes an Improved Tunicate Swarm Algorithm with Deep Learning for Osteosarcoma Detection and Classification (ITSA-DLODC) approach on pathological imageries. The proposed ITSA-DLODC method mainly enhances the recognition and classification of osteosarcoma on HSI. To attain this, the presented ITSA-DLODC method performs feature extraction using ShuffleNet convolutional neural network model. Besides, the ITSA-based hyperparameter optimizer is exploited to finetune the hyperparameters of the ShuffleNet model. Moreover, the salp swarm algorithm (SSA) with convolutional autoencoder (CAE) approach was utilized for the recognition and identification of osteosarcoma. A wide range of analyses can be applied to exemplify the higher performance of the ITSA-DLODC methodology. The simulation study demonstrated the development of the ITSA-DLODC methodology over other present models

Keywords: Intelligent data processing; Data mining; Osteosarcoma; Histopathological images; Computer-aided diagnosis; Deep learning; Tunicate swarm algorithm

1. Introduction

The term Osteosarcoma or Osteogenic Sarcoma can be defined as one type of bone tumor that frequently grows in long bones of arms and legs [1]. But it could cultivate in any bone from the bone cells and rarely affects the soft tissues presented nearby the bones and also affect all age groups, but the number of infected cases in young adults or teens is higher than the other age groups [2]. Osteosarcoma is a general procedure of malignant cancer, which happens in young adults, and the average age for this disease detection is 15 years. The diagnosis of Osteosarcoma patients is still worst because of the metastatic and aggressive behaviour of cancer [3]. Earlier researchers have exposed that the major predictors of diagnosis are a response to chemotherapy, age, tumor dimension, the presence of metastasis at prognosis, gender and contribution of the proximal extremity [4]. Osteogenic Sarcoma can spread to other parts, primarily to the lungs, and rarely to lymph nodes or bone. Metastases are identified in about 15% to 20% of newly identified Osteogenic Sarcoma patients, and the 5yrs survival rate for the patient with metastasis was just 20%. Thus, initial diagnosis at increased risk of metastases and prompt treatment are significant to minimalize the death rate and increase the patient's survival rate [5].

Doi: https://doi.org/10.54216/IJAACI.050104
Received: June 07, 2023 Revised: November 01, 2023 Accepted: December 09, 2023

The task of Osteogenic Sarcoma identification necessitates a lot more professionals and has a time-taking and tedious task if done manually [6]. With the arrival of faster processing systems and innovative technology, the medical sector has swapped to automatic detection systems that could forecast the probability of a tumor being malignant or benign with some certainty [7]. The research community illustrated the growth of computer-assisted diagnosis (CAD) systems as means for identifying and segmenting osteosarcoma utilizing different images, which includes magnetic resonance imaging (MRI) examinations and computed tomography (CT). MRI and CT scans, conversely, have disadvantages. So, the researchers for detecting Osteogenic Sarcoma more precisely than MRIs and CT scans [8] use whole slide images (WSIs). The most popular techniques for detecting such cells are by staining tissue samples of infected parts with eosin and hematoxylin. A diagnostician inspects the samples under a microscope after being stained. WSI of good quality has been utilized in this study [9]. Currently, many machine learning (ML)-related researchers were conducted for cancer prognosis, prediction, or evaluating treatment response [10]. But some research studies are related to the analytical performance of many ML with deep learning (DL) techniques in evaluating lung metastases in OS patients.

The study proposes an Improved Tunicate Swarm Algorithm with Deep Learning for Osteosarcoma Detection and Classification (ITSA-DLODC) model on pathological images. The presented ITSA-DLODC technique performs feature extraction using ShuffleNet convolutional neural network (CNN) system. Besides, the ITSA-based hyperparameter optimizer is exploited to finetune the hyperparameters of the ShuffleNet model. Moreover, the salp swarm algorithm (SSA) with convolutional autoencoder (CAE) approach could be utilized in the recognition and identification of osteosarcoma. A widespread range of experimentations is applied to illustrate the greater performance of the ITSA-DLODC methodology.

2. Related Works

Pan et al. [11] introduced the combination of noise decline convolutional AE and feature cross-fusion learning (NRCA-FCFL), which is a common transformer image classification structure to categorize osteosarcoma histologic images. Noise reduction convolution AE could effectively denoise histologic imagery of osteosarcoma, leading to more pure imageries for classifying osteosarcoma. Vaiyapuri et al. [12] presented an HBODL-AOC method (Honey Badger Optimization with DL-related Automatic Osteosarcoma Classification). In this work, image pre-processing is effectuated primarily through contrast-enhancing systems. For extracting features, the above-mentioned system uses a deep CNN (DCNN)-based MobileNet method that includes Adam optimization for hyper-parameter tuning. Lastly, the adaptive neuro-fuzzy inference system (ANFIS) was applied for the HBO system in tune membership function (MF) and osteosarcoma classification.

Gou et al. [13] constructed a classification assistance technique (OHIcsA) related to a generative adversarial network (GAN) and active learning (AL). To train the classifier, a small, labelled training set has been initially used by the system. After, the useful instances from unlabeled images were chosen for practised annotation. In [14], transfer learning (TL) methods, pre-trained CNN, have been implemented to public data on osteosarcoma histological imageries for identifying necrotic imageries. Next, to enrich the accuracy of the outputs, TL methods, i.e., VGG19 and InceptionV3, are trained and used on WSIs without patches. Lastly, the methods were implemented for different classifier issues, i.e., multi-class and binary classifiers. In [15], an active detection approach was formulated through the Fractional-Harris Hawks Optimizer-based GAN (F-HHO-based GAN) to recognize osteosarcoma at an initial stage. In this study, the F-HHO was modelled by compiling HHO and Fractional Calculus.

In [16], a deep architecture with Siamese network (DS-Net) has been developed. The DS-Net, based on full convolutional networks, is made up of a classification network and an auxiliary supervision network (ASN). The structure of ASN relies upon the network of Siamese intends to overcome the issue of a smaller trained set (the key blockage of DL in healthcare image). The classifier network leverages the attributes derived by the ASN to do precise classification. In [17], the author implemented the DL method to classify osteosarcoma cells. Osteosarcoma has a typical bone tumor happening mostly in young adults or children. Images of these examples were recorded by an optical microscope.

3. The Proposed Model

In this paper, an innovative ITSA-DLODC technique for automatic osteosarcoma detection on histopathology images. The proposed ITSA-DLODC technique improved the identification and classification of osteosarcoma on HI. It follows a sequence of processes, i.e. the ShuffleNet feature extractor, ITSA-based hyperparameter tuning, CAE-based classification, and SSA-related parameter optimization. Fig. 1 shows the work flow of the ITSA-DLODC algorithm.

Doi: https://doi.org/10.54216/IJAACI.050104
Received: June 07, 2023 Revised: November 01, 2023 Accepted: December 09, 2023

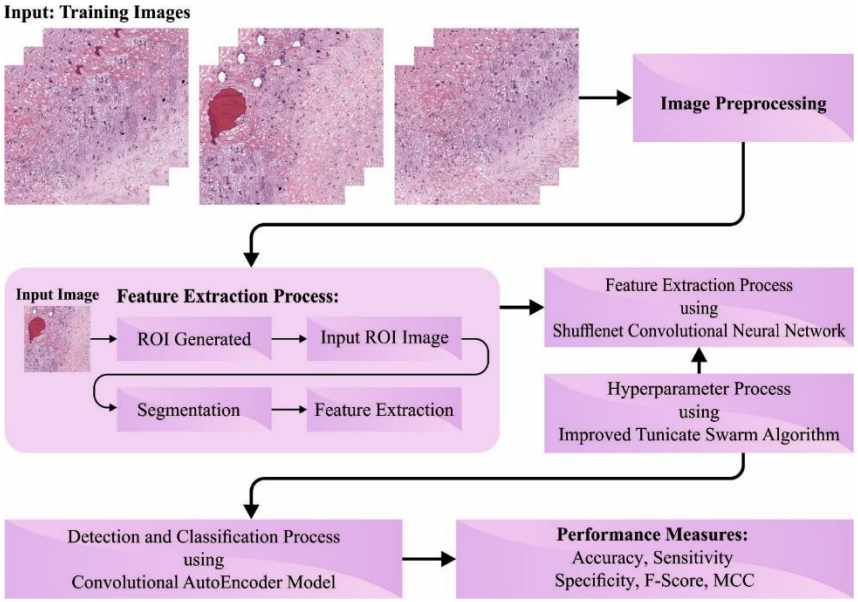

Figure 1: Workflow of ITSA-DLODC algorithm

A. Feature Extraction

Primarily, the ShuffleNet architecture has been employed for producing optimal feature vectors. In this work, we applied the ShuffleNet, an effective DL architecture that was constructed using mobile devices [18]. The study applied the shufflenetv1 version of the pre-trained ShuffleNet architecture based on the computational resource (hardware), to attain improved performance at a lesser computation cost. The presented method is deeper than typical CNN using 50 learnable layers, viz, 48 groups and 1 convolution layer followed by the fully connected (FC) layer. The model has an overall 172 layers, together with 33 Rectified Linear Unit (ReLU) layers, 1 Maxpooling layer, 49 Batch normalization (BN) layers, classification, layer 4 average pooling layers, and softmax layer. The structure applies 4 pooling layers for reducing the total computation struggle.

The first layer is the input, which accepts input imageries of 224×224 size for processing. The initial convolution layer extract features from 224×224 images with 24 filters (kernels) of 3x3 size using the stride of 2x2 for making the mapping feature. The output of mapping features (convolution layers) is evaluated by:

$$s(i,j) = (I \times K)(i,j) = \sum_{n}\sum_{m} I\,(m,n)K(i-m,j-n) \tag{1}$$

In Eq. (1), s characterizes the resultant mapping feature, K denotes the kernel of the existing convolutional layer and i represent the input image. When executing a convolutional function on the input images, the resultant of size $0 = ((i-k)+2p)/(s+1)$ was generated, whereas i characterize input, p means padding, k shows kernel size and s denotes steps.

The ShuffleNet module using stride of 2x2 obtains the resultant mapping feature of the preliminary convolution layer. The ShuffleNet unit contains 3 convolutional functions, viz., 3x3 depthwise convolutions and two 1x1 pointwise group convolutions. The initial pointwise group convolutional layer is monitored by channel shuffle process, BN and ReLu activation function. Relu activation is applied since it is direct and effective.

$$f(x) = \begin{cases} 0, x < 0 \\ x, x \leq 0 \end{cases} \tag{2}$$

Using positive value, the ReLU activate neurons while using negative values, the ReLU deactivate neurons (set neurons to zero). The 2nd and 3rd Conv operations, viz., 3x3 depthwise and 1x1 pointwise convolutional layer, are followed by BN. Similarly, this layer completes feature extractor. Softmax activation determines the classification possibility through the previous layer. The working process of FC layer was revealed in Eq. (3).

$$a_i = \sum_{j=0}^{m\times n-1} w_{ij} \times x_i + b_i \tag{3}$$

Doi: https://doi.org/10.54216/IJAACI.050104
Received: June 07, 2023 Revised: November 01, 2023 Accepted: December 09, 2023

In Eq. (13), $m, n, i\ and\ d$ represents the width, height, index and depth of FC layer output, respectively. Moreover, w denotes the weights and b refers to the bias.

B. Hyperparameter Tuning using ITSA

In this work, the ITSA optimally picks the hyperparameter rate of the ShuffleNet model. TSA is a biologically inspired metamorphic optimization algorithm. In the swarm movement of the Mytilus, this algorithm inspires them to efficiently survive in highly difficult situations of the ocean [19]. Tunicate has a great talent to define the position of food sources in the sea. There exist two characteristics of tunicates like swarm performance and jet momentum used to search the food source of the sea. Tunicate define the optimal food source based on the abovementioned characteristics. Here, the TSA can be enhanced by the mutation and crossover operators hence it is called an improved TSA method which satisfies the load demand of the method. The step-wise procedure can be given as follows.

Step1: Initialized

TSA input vector gets initiated.

Ste2: Random generation

The spontaneous function generates numbers. The lower limits are selected with higher limits.

Step3: Fitness Assessment

Determine the fitness of all the search agents. The fitness calculation is defined by the objective function,

$$Obj\ =\ Min\ (voltage\ deviation, harmonics) \tag{4}$$

Step4: Position Upgrading

Based on fitness, the mass behavior of the Mytilus and jet velocity are defined and position upgrading based on Eq. (5),

$$t_p = \left(\frac{\rho}{y} + 1\right) = \frac{\bar{t_p}(y) + t_p\left(\frac{\rho}{y} + 1\right)}{2 + 0_1}, \tag{5}$$

Here $\tilde{\square}_{\square}(y)$ denotes the location of the tunicate, and 01 shows the random integer.

Step5: Crossover along mutation:

By rearranging the Mytilus swarm place, the upgrading function exploits the shortcut with the mutation operator. Both individuals generate a new solution package; the shortcut ratio can be accomplished.

$$X_{over} = \frac{\delta}{\kappa}, \tag{6}$$

In Eq. (6), κ implies the distance of individuals and δ denotes the count of an individual's crossover.

$$Y_{mu} = \frac{\pi}{\mu} \tag{7}$$

In Eq. (7), π shows transmit point, $and\ L$ represents the distance of individuals.

Step6: Boundary analysis

Verify if the renewal searching agent position is inside and outside the limits.

Step7: Calculation fitness

Defines the position of Mytilus FF with error operation resolution

Step8: End

When the termination condition was satisfied instead of the optimum solution for searching, return to Step 4.

C. Image Classification using Optimal CAE Model

Doi: https://doi.org/10.54216/IJAACI.050104
Received: June 07, 2023 Revised: November 01, 2023 Accepted: December 09, 2023

The CAE approach was exploited for automated osteosarcoma detection and classification, The basic principle of AE is that a single hidden layer has fewer nodes when compared to other layers in the model and performs as a bottleneck [20]. This reduced hidden layer characterizes the relevant features of an image with limited number of data. Using image input, AE converts the unstructured data as a feature vector that is handled by other ML techniques. Fig. 2 signifies the structure of CAE.

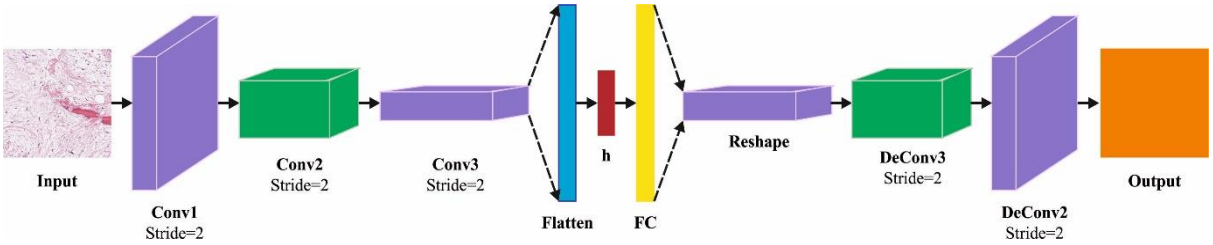

Figure 2: Architecture of CAE

The CAE is classified into the two major stages of interpreting and encoding. It given below:

$$O_m(i,j) = a\left(\sum_{d=1}^{D}\sum_{u=-2k-1}^{2k+1}\sum_{v=-2k-1}^{2k+1} F_{m_d}^{(1)}(u,v) I_d(i-u, j-v)\right) \quad (8)$$

$$m = 1, \cdots, n$$

In Eq. (8), $F \in \{F_1^{(1)}, F_2^{(1)}, \ldots, F_n^{(1)},\}$ refer to a convolutional filter; with convolution amongst inputs as $I = \{I_1, \ldots, I_D\}$ represent the input by integrating nonlinear functions:

$$z_m = O_m = a\left(I * F_m^{(1)} + b_m^{(1)}\right) m = 1, \ldots, m \quad (9)$$

In Eq. (9), $b_m^{(1)}$ denotes the bias and the count of zeros it is needed to pad the input: $dim\,(I) = dim\,(decode(encode(I)))$. Lastly, the encoded convolution is equivalent to:

$$O_w = O_h = (I_w + 2(2k+1) - 2) - (2k+1) + 1 \quad (10)$$

$$= I_w + (2k+1) - 1$$

The decoded convolutional stage generates n feature maps $z_{m=1,\ldots n}$. The reconstructed outcome I am the result of convolutional among the number of mapping features $Z = \{z_{i=1}\}^n$ and number of convolution filtering $F^{(2)}$.

$$\tilde{I} = a\left(Z * F_m^{(2)} + b^{(2)}\right) \quad (11)$$

$$O_w = O_h = (I_w + (2k+1) - 1) - (2k+1) + 1 = I_w = I_h \quad (12)$$

Eq. (12) demonstrates the decoded convolution with I dimensions. The input dimension is equivalent to the output dimension.

SSA is a SI optimizer approach that act out the effort behaviors of the salp populace chain in the sea and was proposed in 2017 [21]. Lastly, the SSA optimally selects the hyperparameter values of the SSA. The SSA comes under the type of swarm-based algorithm with the meta-heuristic method. Also, SSA simulates a salp swarming while finding food in the ocean. salp forms shoals On the ocean floor named salp chain. In this work, the salp leader is at the front of the chain and the remaining salp is named the followers. Like other swarm-based methods, salp position can be determined in an s-dimensional searching space, whereas s denotes the amount of parameters in the presented issue. Thus, the position of each salp is stored in a 2D matrix named χ. Also, it is considered that a food source named P from the searching space as aim of swarms. Simulation using $2D$ space shows that the model displays the behavior of the SSA in n-dimensional space.

As demonstrated in the mathematical Eq. (13) a set of X salps out of n is characterized as a 2D matrix. The food source that which salps chain reaches is represented by F in the search space.

Doi: https://doi.org/10.54216/IJAACI.050104
Received: June 07, 2023 Revised: November 01, 2023 Accepted: December 09, 2023

$$X_i = \begin{bmatrix} x_1^1 & x_2^1 & \dots & x_d^1 \\ x_1^2 & x_2^2 & & x_d \\ \vdots & & \ddots & \vdots \\ x_1^n & x_2^n & \dots & x_d^n \end{bmatrix} \quad (13)$$

The mathematical expression for SSA is given based on the SSA model and the leader location equation is given below:

$$x_j^1 = \begin{cases} F_j + c_1 + ((ub_j - lb_j)c_2 + lb_j)c_3 \geq 0 \\ F_j - c_1 + ((ub_j - lb_j)c_2 + lb_j)c_3 < 0 \end{cases} \quad (14)$$

The food source location at $the\ j$ dimension can be specified as F_j. ub_j and lb_j upper and lower boundaries of $j\text{-}th$ dimensions. The random number is allotted to c_1, c_2 and c_3. The coefficient parameter c_1 is significant in SSA since it gives a balance among exploration as well as exploitation abilities.

$$c_1 = 2e^{-\left(\frac{4l}{L}\right)^2} \quad (15)$$

The present iteration is named l and the higher iteration count is L. The random number is uniformly generated within [1,0] are variables c_2 and c_3. Here, this parameter determines the next location in the $j\text{-}th$ dimension resulting in negative or positive infinity and the step size. The follower position is upgraded by the subsequent equations (Newton's law of motion):

$$\chi_j^i = \frac{1}{2}at^2 + v_0^t \quad (16)$$

When $i \geq 2$, t shows the time, x_j^i represents the salp location of the $i\text{-}th$ followers at $j\text{-}th$ variable, v_0 shows the initial speed and $a = \frac{v_{final}}{v_0}$ dimana $v = \frac{x - x_0}{t}$. Since iteration indicates a time in the optimizer, the conflict among iterations is 1, afterwards assuming $v_0 = 0$, this formula is expressed as:

$$x_j^i = \frac{1}{2}\left(x_j^i + x_j^{i-1}\right) \quad (17)$$

Where x_j^i denotes the position of $i\text{-}th$ followers at $the\ j\text{-}th$ dimension.

Algorithm 1: Pseudocode of SSA

Input parameters: Population Size, Iteration count, Min and Max Values

Initializing the salp populace $x_i(i = 1,2,3, \dots, n)$ assuming ub and lb

While (ending criteria is not met) do

Evaluate the fitness of every searching agent (salp)

$F = the$ optimum searching agent

Upgrade c_1 by Eq. (15)

    For every salp (x)

        If ($i == 1$)

            Upgrading the leading salp place using Eq. (14)

        Else

            Upgrading the follower salp place using Eq. (17)

        End

    End

Adjust the salps dependent upon the lower and upper boundaries of variable

Doi: https://doi.org/10.54216/IJAACI.050104
Received: June 07, 2023 Revised: November 01, 2023 Accepted: December 09, 2023

End

return F

Output: Global optimum solution

The salp chain is simulated by Eqs. (14) and (17). F_j shows the location of the food source at $the\ j\text{-}th$ dimension, u_{bj} and l_{bj} indicate the upper and lower limits in the searching space at $the\ j\text{-}th$ dimensional, and c_1, c_2 and c_3 represent a random variable that is uniformly generated from the interval. The SSA pseudocode was projected in Algorithm 1. The SSA manner grows the FF to obtain higher classifier accuracy. And also determines a positive number to signify the superior effectiveness of candidate performances. The decrease of the rate of classifier error is supposed as FF.

$$fitness(x_i) = ClassifierErrorRate(x_i)$$
$$= \frac{number\ of\ misclassified\ samples}{Total\ number\ of\ samples} * 100 \quad (18)$$

4. Experimental Validation

The simulation outcome of the ITSA-DLODC system was surveyed on the database [22], comprising 1144 samples with a 3 classes as given in Table 1. The database includes 1144 images with resolution of 1024x1024, with the subsequent distribution: 263(23%) necrotic tumor images, 345(30%) viable tumor tiles, and 536(47%) non-tumor images. Fig. 3 illustrates the sample imageries.

Table 1: Details on database

Classes	No. of Samples
Non-Tumor	536
Necrotic Tumor	263
Viable Tumor	345
Total Number of Samples	1144

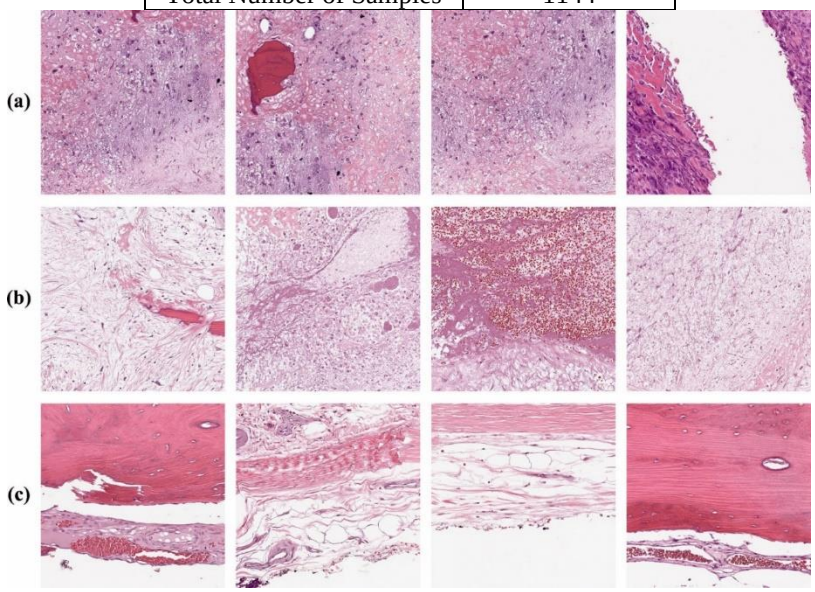

Figure 3: a) Viable-Tumor b) Necrotic-Tumor c) Non-Tumor

In Fig. 4, the confusion matrices of the ITSA-DLODC model at osteosarcoma classification can be demonstrated. The outcomes highlighted the ITSA-DLODC technique gains effectual identification of three types of samples.

In Table 2 and Fig. 5, a detailed result of the ITSA-DLODC method with 80:20 of the TRS/TSS can be illustrated. The experimental outcome indicates improved outcomes below all classes. For example, with 80% of TRS, the ITSA-DLODC methodology offers an average $accu_y$ of 97.81%, $sens_y$ of 96.88%, $spec_y$ of 98.36%, F_{score} of

Doi: https://doi.org/10.54216/IJAACI.050104
Received: June 07, 2023 Revised: November 01, 2023 Accepted: December 09, 2023

96.69%, and MCC of 94.99%. Furthermore, with 20% of TSS, the ITSA-DLODC system offers average $accu_y$ of 99.42%, $sens_y$ of 99.22%, $spec_y$ of 99.60%, F_{score} of 99.06%, and MCC of 98.63%.

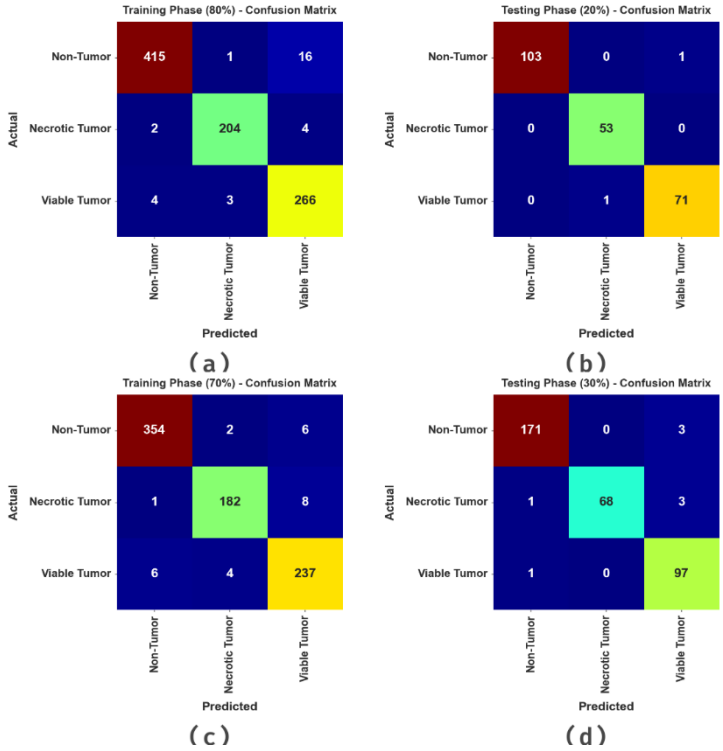

Figure 4: Confusion matrices of ITSA-DLODC algorithm (a-b) TRS/TSS of 80:20 and (c-d) TRS/TSS of 70:30

Table 2: Classification outcome of ITSA-DLODC model with 80:20 of TRS/TSS

Class	$Accu_y$	$Sens_y$	$Spec_y$	F_{Score}	MCC
TRS (80%)					
Non-Tumor	97.49	96.06	98.76	97.30	94.98
Necrotic Tumor	98.91	97.14	99.43	97.61	96.90
Viable Tumor	97.05	97.44	96.88	95.17	93.10
Average	97.81	96.88	98.36	96.69	94.99
TSS (20%)					
Non-Tumor	99.56	99.04	100.00	99.52	99.12
Necrotic Tumor	99.56	100.00	99.43	99.07	98.79
Viable Tumor	99.13	98.61	99.36	98.61	97.97
Average	99.42	99.22	99.60	99.06	98.63

Doi: https://doi.org/10.54216/IJAACI.050104
Received: June 07, 2023 Revised: November 01, 2023 Accepted: December 09, 2023

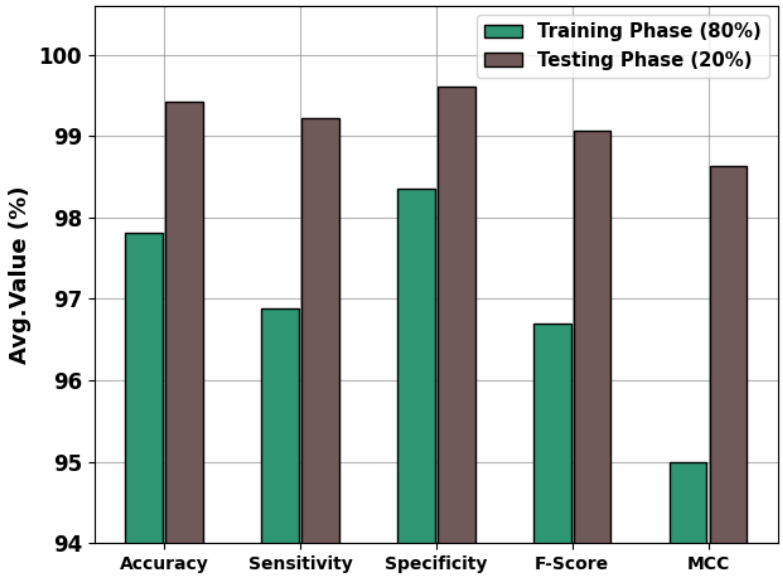

Figure 5: Average outcome of ITSA-DLODC algorithm with 80:20 of TRS/TSS

Fig. 6 observes the $accu_y$ of the ITSA-DLODC approach in the training and validation model in 80:20 of the TRS/TSS. The achieved result shows that the ITSA-DLODC method accomplishes enhancing $accu_y$ values over higher epochs. Furthermore, the enhancing validation $accu_y$ over training $accu_y$ exposes that the ITSA-DLODC approach attains efficiently on 80:20 of TRS/TSS.

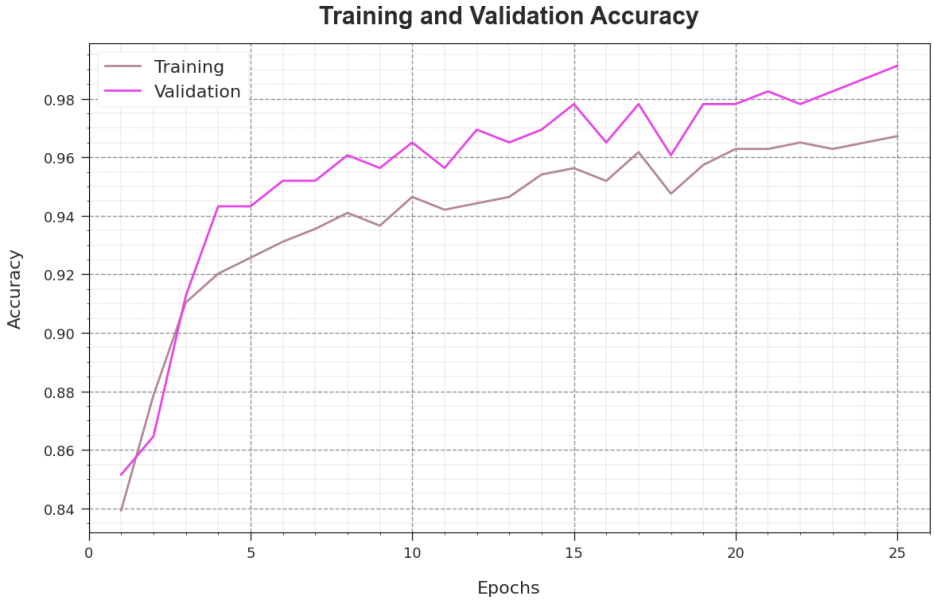

Figure 6: Accuracy curve of ITSA-DLODC algorithm

The loss study of the ITSA-DLODC methodology in training and validation is defined in Fig. 7. The result signified that the ITSA-DLODC methodology obtains neighboring values of training and validation loss. It must be stated that the ITSA-DLODC method attains efficiently.

A complete precision-recall (PR) outcome of the ITSA-DLODC system can be established in Fig. 8. The obtained outcomes exhibits as the ITSA-DLODC algorithm results in maximal values of PR. Moreover, it could be clear that the ITSA-DLODC model can achieve higher PR values in each class.

Doi: https://doi.org/10.54216/IJAACI.050104
Received: June 07, 2023 Revised: November 01, 2023 Accepted: December 09, 2023

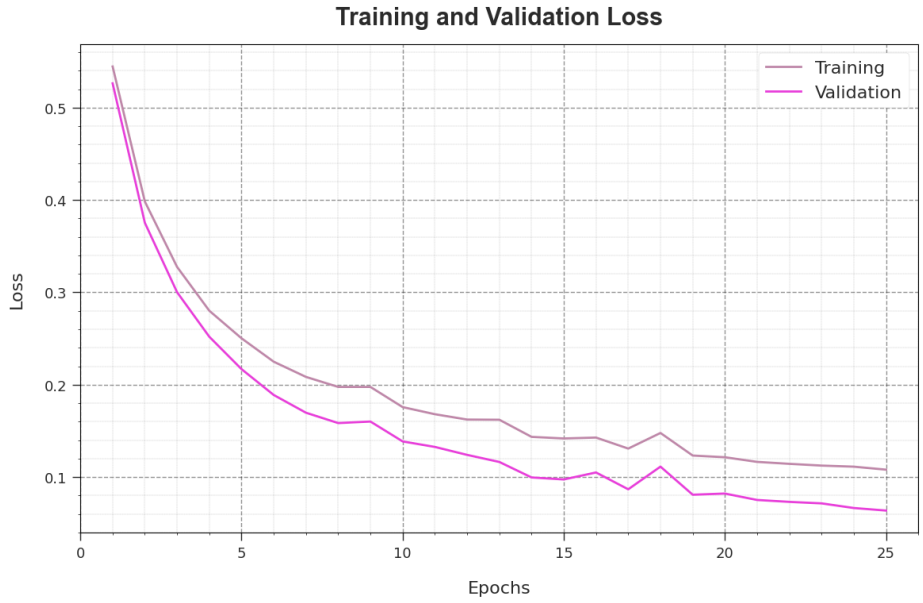

Figure 7: Loss curve of ITSA-DLODC algorithm

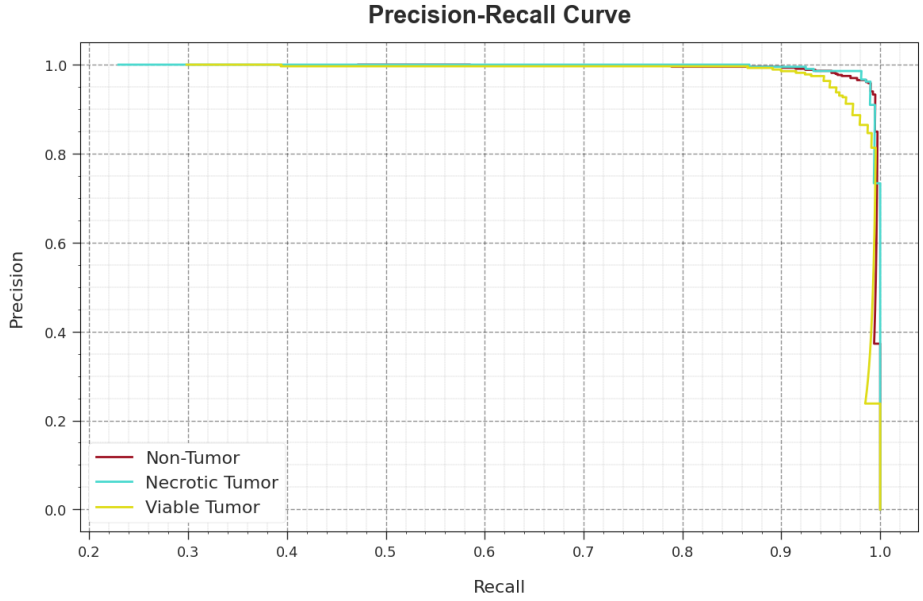

Figure 8: PR analysis of ITSA-DLODC methodology

In Fig. 9, a ROC examination of the ITSA-DLODC approach has been exposed. The outcome described that the ITSA-DLODC system resulted in higher ROC values. Moreover, it can be apparent that the ITSA-DLODC algorithm can range higher ROC values in every class.

Doi: https://doi.org/10.54216/IJAACI.050104
Received: June 07, 2023 Revised: November 01, 2023 Accepted: December 09, 2023

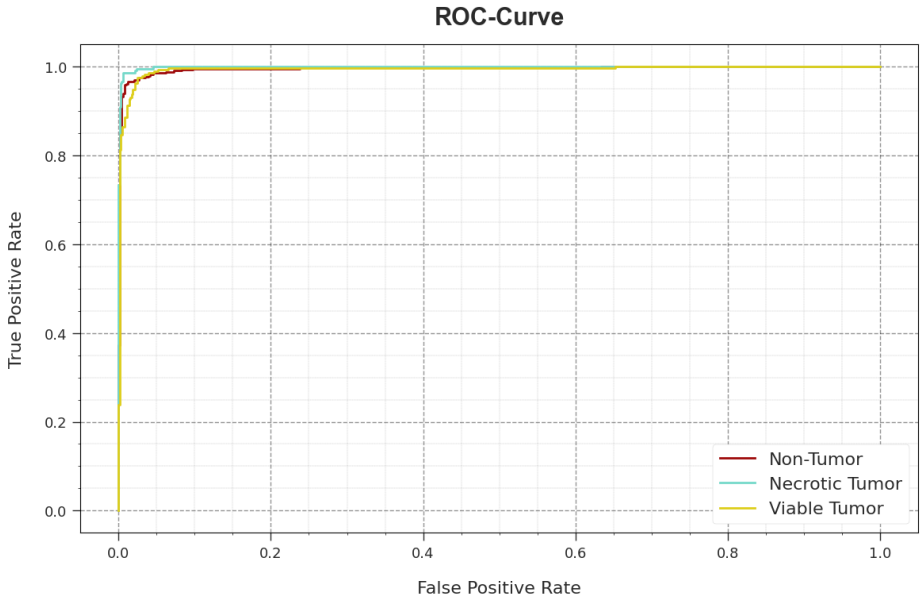

Figure 9: ROC curve of ITSA-DLODC algorithm

In Table 3 and Fig. 10, a complete outcome of the ITSA-DLODC approach is exemplified. The experimental outcome represented enhanced performance in every class. For instance, with 70% of TRS, the ITSA-DLODC methodology offers an average $accu_y$ of 97.75%, $sens_y$ of 96.34%, $spec_y$ of 98.29%, F_{score} of 96.38%, and MCC of 94.68%. Additionally, with 30% of TSS, the ITSA-DLODC methodology gives average $accu_y$ of 98.45%, $sens_y$ of 97.23%, $spec_y$ of 98.79%, F_{score} of 97.41%, and MCC of 96.24%.

Table 3: Classification outcome of ITSA-DLODC algorithm with 70:30 of TRS/TSS

Class	$Accu_y$	$Sens_y$	$Spec_y$	F_{Score}	MCC
TRS (70%)					
Non-Tumor	98.12	97.79	98.40	97.93	96.22
Necrotic Tumor	98.12	95.29	99.01	96.04	94.82
Viable Tumor	97.00	95.95	97.47	95.18	93.01
Average	97.75	96.34	98.29	96.38	94.68
TSS (30%)					
Non-Tumor	98.55	98.28	98.82	98.56	97.09
Necrotic Tumor	98.84	94.44	100.00	97.14	96.48
Viable Tumor	97.97	98.98	97.56	96.52	95.14
Average	98.45	97.23	98.79	97.41	96.24

Fig. 11 examines the $accu_y$ of the ITSA-DLODC model in training and validation model at 70:30 of the TRS/TSS. The outcome concluded that the ITSA-DLODC approach gains maximal $accu_y$ values over greater epochs. Followed by, the improving validation $accu_y$ over training $accu_y$ shows that the ITSA-DLODC system obtains effectively under 70:30 of TRS/TSS.

The loss result of the ITSA-DLODC system in training and validation is exposed in 70:30 of the TRS/TSS in Fig. 12. The result referred that the ITSA-DLODC algorithm obtains nearby values of training and validation loss. Note that the ITSA-DLODC approach reaches effectually in 70:30 of the TRS/TSS.

Doi: https://doi.org/10.54216/IJAACI.050104
Received: June 07, 2023 Revised: November 01, 2023 Accepted: December 09, 2023

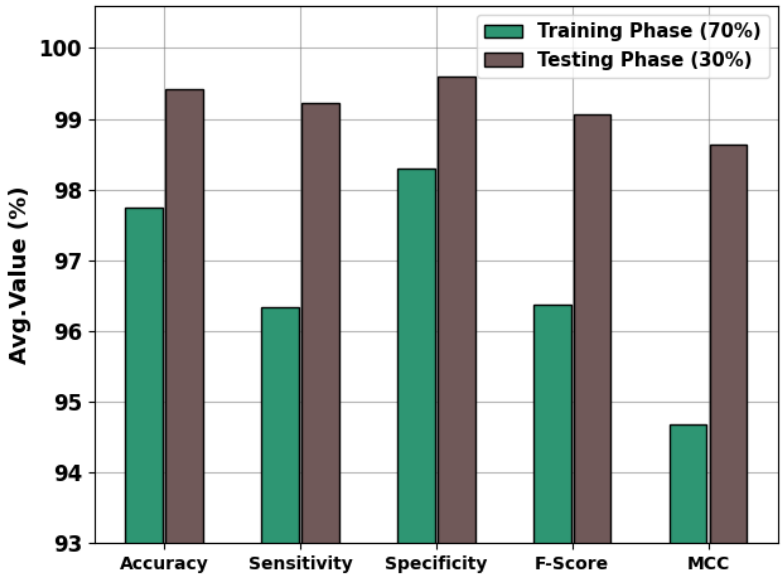

Figure 10: Average outcome of ITSA-DLODC algorithm with 70:30 of TRS/TSS

Figure 11: Accuracy curve of ITSA-DLODC system on 70:30 of TRS/TSS

A comprehensive PR analysis of the ITSA-DLODC model is well-known on 70:30 of the TRS/TSS in Fig. 13. The result stated that the ITSA-DLODC technique results in highest values of PR. Besides, it is noticeable that the ITSA-DLODC model gain greater PR values in every class.

Doi: https://doi.org/10.54216/IJAACI.050104
Received: June 07, 2023 Revised: November 01, 2023 Accepted: December 09, 2023

Figure 12: Loss curve of ITSA-DLODC model on 70:30 of TRS/TSS

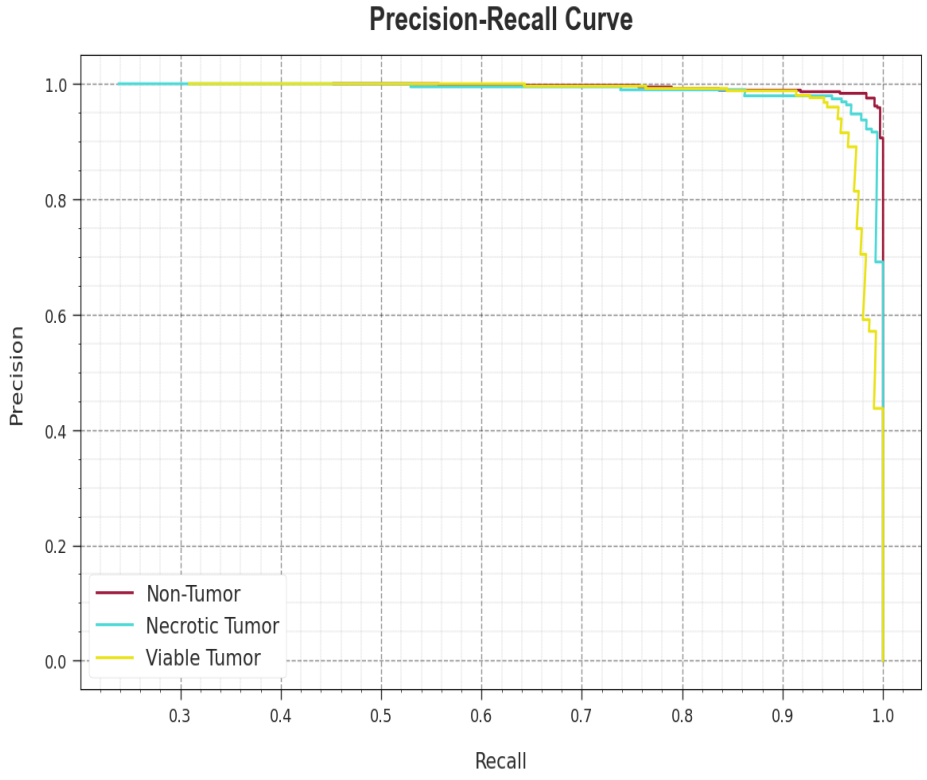

Figure 13: PR curve of ITSA-DLODC model on 70:30 of TRS/TSS

In Fig. 14, a ROC study of the ITSA-DLODC procedure is visible at 70:30 of the TRS/TSS. The result definite that the ITSA-DLODC system ensued in better ROC values. Furthermore, it is clear that the ITSA-DLODC approach can range higher ROC values on 3 classes.

Doi: https://doi.org/10.54216/IJAACI.050104
Received: June 07, 2023 Revised: November 01, 2023 Accepted: December 09, 2023

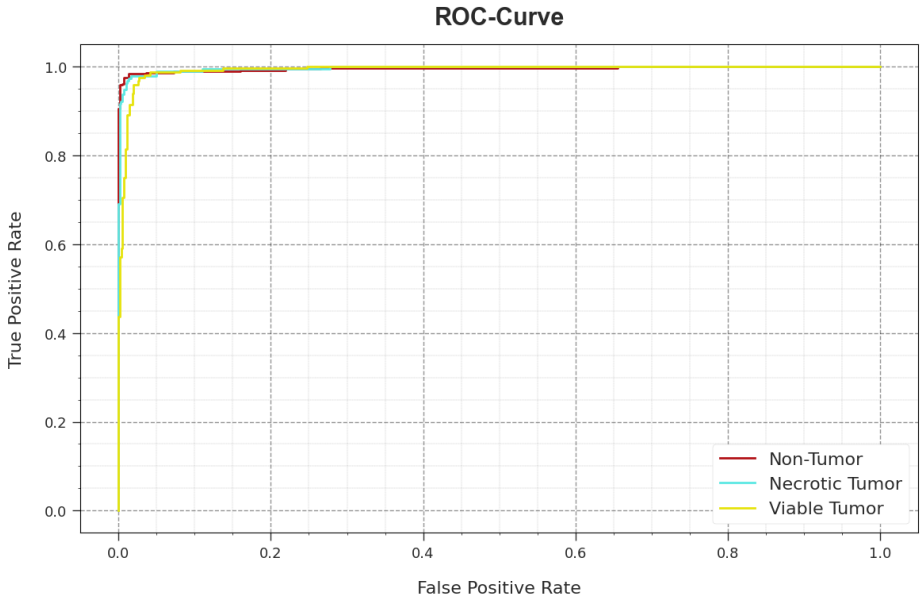

Figure 14: ROC curve of ITSA-DLODC methodology on 70:30 of TRS/TSS

A detailed comparison outcome of the ITSA-DLODC model with other techniques is shown in Table 4 and Fig. 15 [3]. The simulation outcome highlighted that the CNN and SVM model gains poor results over existing techniques while the VGGNet model reaches slightly increased outcomes. Simultaneously, the K-means and VGG19 model offers moderately closer results. Meanwhile, the NRCA-FCFL and SGDM models accomplish reasonable performance with closer $accu_y$ of 99.01% and 99.10% correspondingly.

Table 4: Comparative outcome of ITSA-DLODC algorithm with existing methodologies

Methods	$Accu_y$	$Sens_y$	$Spec_y$	F_{Score}
ITSA-DLODC	99.42	99.22	99.60	99.06
CNN	84.00	95.51	98.81	95.66
K-means	95.50	97.00	95.72	97.79
VGGNet	92.00	96.63	97.28	96.97
VGG19	96.00	96.82	95.80	98.31
SVM	89.90	97.08	95.03	95.31
NRCA-FCFL	99.01	97.22	96.18	96.38
SGDM	99.10	97.90	95.59	95.16

Doi: https://doi.org/10.54216/IJAACI.050104
Received: June 07, 2023 Revised: November 01, 2023 Accepted: December 09, 2023

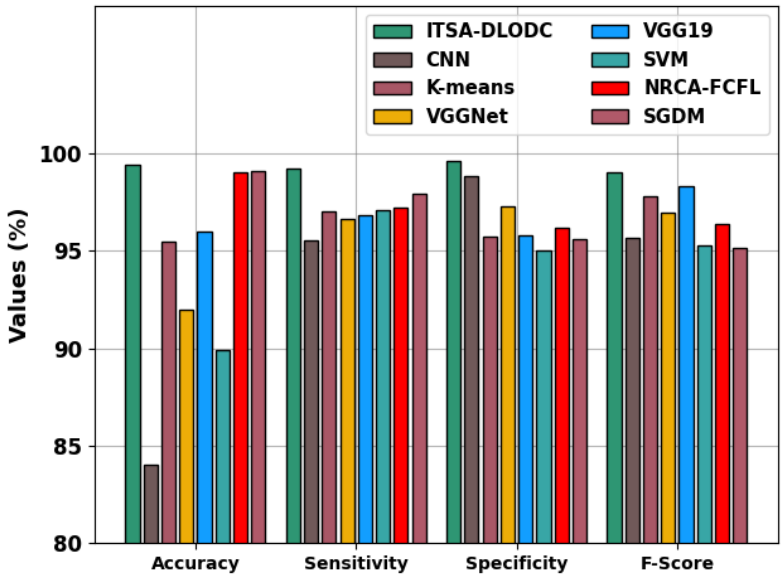

Figure 15: Comparative outcome of ITSA-DLODC approach with recent methodologies

However, the ITSA-DLODC technique exhibits maximum performance with $accu_y$ of 99.42%, $sens_y$ of 99.22%, $spec_y$ of 99.60%, and F_{score} of 99.06%. Therefore, the ITSA-DLODC technique shows promising performance over other recent DL approaches.

5. Conclusion

In this paper, a novel ITSA-DLODC approach was presented for automatic osteosarcoma recognition on HSI. The proposed ITSA-DLODC technique improved the recognition and classification of osteosarcoma on histopathological images. It follows a sequence of processes namely ShuffleNet feature extractor, ITSA-based hyperparameter tuning, CAE-based identification, and SSA-based parameter optimizer. To achieve this, the presented ITSA-DLODC approach completes feature extraction utilizing ShuffleNet convolutional neural network model. Besides, the ITSA-based hyperparameter optimizer is exploited to finetune the hyperparameters of the ShuffleNet technique. Moreover, SSA with CAE approach was employed for the recognition and identification of osteosarcoma. A huge range of experimentations is completed to exemplify the superior result of the ITSA-DLODC methodology. The simulation analysis demonstrated the improvement of ITSA-DLODC system over other existing techniques.

Funding: "This research received no external funding"

Conflicts of Interest: "The authors declare no conflict of interest."

References

[1] Arunachalam, H.B., Mishra, R., Daescu, O., Cederberg, K., Rakheja, D., Sengupta, A., Leonard, D., Hallac, R. and Leavey, P., 2019. Viable and necrotic tumor assessment from whole slide images of osteosarcoma using machine-learning and deep-learning models. PloS one, 14(4), p.e0210706.

[2] Tang, H., Sun, N. and Shen, S., 2021. Improving generalization of deep learning models for diagnostic pathology by increasing variability in training data: Experiments on osteosarcoma subtypes. Journal of Pathology Informatics, 12(1), p.30.

[3] Nasir, M.U., Khan, S., Mehmood, S., Khan, M.A., Rahman, A.U. and Hwang, S.O., 2022. IoMT-Based Osteosarcoma Cancer Detection in Histopathology Images Using Transfer Learning Empowered with Blockchain, Fog Computing, and Edge Computing. Sensors, 22(14), p.5444.

[4] Patkar, S., Beck, J., Harmon, S., Mazcko, C., Turkbey, B., Choyke, P., Brown, G.T. and LeBlanc, A., 2023. Deep domain adversarial learning for species-agnostic classification of histologic subtypes of osteosarcoma. The American Journal of Pathology, 193(1), pp.60-72.

[5] Mahore, S., Bhole, K. and Rathod, S., 2021, July. Comparative analysis of machine learning algorithm for classification of different osteosarcoma types. In 2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT) (pp. 1-5). IEEE.

Doi: https://doi.org/10.54216/IJAACI.050104
Received: June 07, 2023 Revised: November 01, 2023 Accepted: December 09, 2023

[6] Ho, D.J., Agaram, N.P., Schüffler, P.J., Vanderbilt, C.M., Jean, M.H., Hameed, M.R. and Fuchs, T.J., 2020, October. Deep interactive learning: an efficient labeling approach for deep learning-based osteosarcoma treatment response assessment. In International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 540-549). Springer, Cham.
[7] D'Acunto, M., Martinelli, M. and Moroni, D., 2019. From human mesenchymal stromal cells to osteosarcoma cells classification by deep learning. Journal of Intelligent & Fuzzy Systems, 37(6), pp.7199-7206.
[8] Varalakshmi, P., Priyamvadan, A.V. and Rajakumar, B.R., 2022, January. Predicting Osteosarcoma using eXtreme Gradient Boosting Model. In 2022 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI) (pp. 1-6). IEEE.
[9] Wu, J., Yang, S., Gou, F., Zhou, Z., Xie, P., Xu, N. and Dai, Z., 2022. Intelligent segmentation medical assistance system for mri images of osteosarcoma in developing countries. Computational and Mathematical Methods in Medicine, 2022.
[10] Mahore, S., Bhole, K. and Rathod, S., 2021, August. Machine Learning approach to classify and predict different Osteosarcoma types. In 2021 8th International Conference on Signal Processing and Integrated Networks (SPIN) (pp. 641-645). IEEE.
[11] Pan, L., Wang, H., Wang, L., Ji, B., Liu, M., Chongcheawchamnan, M., Yuan, J. and Peng, S., 2022. Noise-reducing attention cross fusion learning transformer for histological image classification of osteosarcoma. Biomedical Signal Processing and Control, 77, p.103824.
[12] Vaiyapuri, T., Jothi, A., Narayanasamy, K., Kamatchi, K., Kadry, S. and Kim, J., 2022. Design of a Honey Badger Optimization Algorithm with a Deep Transfer Learning-Based Osteosarcoma Classification Model. Cancers, 14(24), p.6066.
[13] Gou, F., Liu, J., Zhu, J. and Wu, J., 2022, November. A Multimodal Auxiliary Classification System for Osteosarcoma Histopathological Images Based on Deep Active Learning. In Healthcare (Vol. 10, No. 11, p. 2189). Multidisciplinary Digital Publishing Institute.
[14] Anisuzzaman, D.M., Barzekar, H., Tong, L., Luo, J. and Yu, Z., 2021. A deep learning study on osteosarcoma detection from histological images. Biomedical Signal Processing and Control, 69, p.102931.
[15] Badashah, S.J., Basha, S.S., Ahamed, S.R. and Subba Rao, S.P.V., 2021. Fractional-Harris hawks optimization-based generative adversarial network for osteosarcoma detection using Renyi entropy-hybrid fusion. International Journal of Intelligent Systems, 36(10), pp.6007-6031.
[16] Fu, Y., Xue, P., Ji, H., Cui, W. and Dong, E., 2020. Deep model with Siamese network for viable and necrotic tumor regions assessment in osteosarcoma. Medical Physics, 47(10), pp.4895-4905.
[17] Charchekhandra, B. (2023). The Reading and Analyzing Of The Brain Electrical Signals To Execute a Control Command and Move an Automatic Arm. *Pure Mathematics for Theoretical Computer Science*, 1(1), 08-16.
[18] Ullah, N., Khan, J.A., El-Sappagh, S., El-Rashidy, N. and Khan, M.S., 2023. A Holistic Approach to Identify and Classify COVID-19 from Chest Radiographs, ECG, and CT-Scan Images Using ShuffleNet Convolutional Neural Network. Diagnostics, 13(1), p.162.
[19] Reddy, C.S.R., Prasanth, B.V. and Chandra, B.M., 2023. Active power management of grid-connected PV-PEV using a Hybrid GRFO-ITSA technique. Science and Technology for Energy Transition, 78, p.7.
[20] Kowsari, K., Sali, R., Ehsan, L., Adorno, W., Ali, A., Moore, S., Amadi, B., Kelly, P., Syed, S. and Brown, D., 2020. Hmic: Hierarchical medical image classification, a deep learning approach. Information, 11(6), p.318.
[21] Kristiyanti, D.A., Sitanggang, I.S. and Nurdiati, S., 2023. Feature Selection Using New Version of V-Shaped Transfer Function for Salp Swarm Algorithm in Sentiment Analysis. Computation, 11(3), p.56.
[22] https://wiki.cancerimagingarchive.net/plugins/servlet/mobile?contentId=52756935#content/view/52756935

Doi: https://doi.org/10.54216/IJAACI.050104
Received: June 07, 2023 Revised: November 01, 2023 Accepted: December 09, 2023